

# Exhibit K

THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF NEW JERSEY  
CAMDEN VICARAGE

IN RE: VALSARTAN PRODUCTS LIABILITY LITIGATION  <i>This document applies to all actions</i>	Master Docket Number 1:19-md-2875 (RBK)(JS)  <b>ORDER</b>
--	--

CASE MANAGEMENT ORDER No. 8  
(ELECTRONIC DISCOVERY PROTOCOL – STIPULATED)

SCHNEIDER, United States Magistrate Judge

The parties, through their counsel, have stipulated and agreed, this Stipulated Order regarding Electronic Discovery Protocol (“ESI Protocol”) shall apply to all cases in this litigation.

**WHEREAS** certain documents and information have been and may be requested, sought, produced, or exhibited by, between and among the parties to the above-captioned actions that involve valsartan, and any future cases in this consolidated litigation; such records and data that constitute electronically stored information (“ESI”) shall be produced by the Parties in accordance with the following protocol:

**I. DEFINITIONS**

**A.** “Electronically stored information,” or “ESI,” as used herein, means and refers to computer generated information or data of any kind, stored in or on any storage media as defined in Federal Rule of Civil Procedure 34. Non limiting examples of ESI include:

- Digital Communications (e.g., e-mail, voice mail, instant messaging, etc.);

- E-Mail Server Stores (e.g., Gmail, Lotus Domino .NSF or Microsoft Exchange .EDB);
- Word Processed Documents (e.g., Word or WordPerfect files and drafts);
- Spreadsheets and tables (e.g., Excel or Lotus 123 worksheets);
- Collaboration software tools (e.g., Microsoft SharePoint, Microsoft Lync, eRoom);
- Document Managements Systems (e.g., Documentum);
- Accounting Application Data (e.g., QuickBooks, Money, Peachtree data);
- Image and Facsimile Files (e.g., .PDF, .TIFF, .JPG, .GIF images);
- Sound Recordings (e.g., .WAV and .MP3 files);
- Video and Animation (e.g., .AVI and .MOV files);
- Databases (e.g., Access, Oracle, SQL Server data, SAP);
- Contact and Relationship Management Data (e.g., Outlook, ACT!);
- Calendar and Diary Application Data (e.g., Outlook PST, blog entries);
- Presentations (e.g., PowerPoint, Corel Presentations);
- Project Management Application Data;
- Task related data (e.g., Task lists from Outlook PST);
- Computer Aided Design/Drawing Files;
- Backup and Archival Files (e.g., Veritas, Zip, .GHO); and
- Apps (data from smartphone applications used in the normal course of business).

This list is neither exhaustive, nor does it represent that such ESI exists within the parties' data sources or that it will be produced. The parties agree to meet and confer regarding any disputes relating to relevancy, proportionality, or production issues before submission of any disputes to the court.

B. **“Native data format”** means and refers to the format of ESI in which it was generated and/or as used by the producing party in the usual course of its business and in its regularly conducted activities.

C. **“Metadata”** means and refers to information about information or data about data, and includes, without limitation: (i) information embedded in or associated with a native file that is not ordinarily viewable or printable from the application that generated, edited, or modified such native file which describes the characteristics, origins, usage and/or validity of the electronic file and/or (ii) information generated automatically by the operation of a computer or other information technology system when a native file is created, modified, transmitted, deleted or otherwise manipulated by a user of such system.

D. **“Documents”** includes writings, drawings, graphs, charts, photographs, sound recordings, images, and other data, data records or data compilations – stored in any medium (including but not limited to physical, electronic, and cloud-based or cloud sourced media) from which information can be obtained.

E. **“Storage Media”** means an object or device, real or virtualized, including but not limited to a disc, tape, computer, thumb drive or USB drive, mobile device, smartphone, iPad, Cloud (public, private or hybrid) or other device, or other similar means of communication, on which data is or was stored.

## II. **SEARCH TERMS FOR ELECTRONIC DOCUMENTS**

The parties agree that they will cooperate in good faith regarding the disclosure and formulation of appropriate search methodology, search terms and protocols, and any TAR/predictive coding prior to using any such technology to narrow the pool of collected

documents to a set to undergo review for possible production. The parties agree to meet and confer as early as possible, to discuss, *inter alia*:

- Search methodology(ies) to be utilized (including but not limited to Boolean searches and technology assisted review/predictive coding);
- Disclosure and selection of all search terms, including foreign language (e.g., Chinese) and search terms based upon actual terminology used in the producing party's documents;
- Search protocol, including but not limited to the custodial and non-custodial sources to which the search methodology(ies) will be applied; and
- Post-search error sampling and sampling/testing reports.

The parties will continue to meet and confer in accordance with the Federal Rules of Civil Procedure and Local Civil Rules of this court regarding any search process issues as necessary and appropriate. Nothing in this protocol, or the subsequent designation of any search terms, shall operate to limit or expand a party's production obligations under applicable law, or limit the ability to agree to meet and confer to modify the search terms, criteria, or methodologies as information is obtained and utilized in the course of the litigation. Nor shall anything in this protocol be deemed to constitute a waiver of any objections a producing party may have to any document request.

This ESI protocol does not address or resolve any other objection to the scope of the parties' respective discovery requests.

### **III. FORMAT OF PRODUCTION**

**A. General Production Formats.** ESI will be produced as single-page TIFF images pursuant to this order except as set forth below. All word processing files produced as TIFF images

shall include all information contained in the original file, including but not limited to tracked changes, comments, and hidden text, except where such information has been redacted for privilege. Where agreed, ESI will be produced in native format with Metadata intact, except where this is not practicable. In this event, the producing party will promptly notify the requesting party of the issue, and the parties will attempt to agree upon an alternative format that is acceptable to the requesting party.

1. Excel (or other spreadsheet formats) will be produced in native format. Native Excel files will be redacted by overwriting the data contained in a particular cell, row, column or tab with the word "REDACTED." These documents will be identified as redacted within the "REDACTED" field in the .DAT. If the producing party deems it necessary to produce redacted Excel (or other spreadsheet formats) in any format other than Native, the parties shall meet and confer before such a production is made.
2. PowerPoint (or other presentation formats) will be produced in native format. Native PowerPoint presentations that require redactions shall be produced as single-page TIFF images, which shall include speaker notes and "hidden" slides as extracted by the software used to process the documents. Color PowerPoint presentations requiring redactions shall be converted to color images as set forth in Section D, and black and white PowerPoint presentations requiring redactions shall be converted to black and white TIFF images, provided that proper grayscale printing is enabled to ensure that any dark colored text is not hidden by dark objects/drawings around the text. If the PowerPoint or slide program contains video or audio

components, the video or audio will be produced as native files with the appropriate attachment relationships.

3. Audio and video files will be produced in native format. Audio and video files will be edited if redactions are required, subject to appropriate identification of any such audio or video files having been redacted and the locations of such redactions.
4. PDF files shall be produced natively with extracted text files. Native PDF files which are redacted shall be redacted by overwriting the privileged data with a delineated box and the word "REDACTED." PDF files with redactions will be identified as redacted within the "REDACTED" field in the .DAT. If the file contains video or audio components, the video or audio will be produced as native files with appropriate links to the parent document.
5. Documents that meet a to-be-determined threshold of foreign language content (*e.g.*, non-English character set) shall be produced natively and as single-page TIFF images, provided, however, that any such documents requiring redaction shall be produced as single-page TIFF images only. The parties agree to meet and confer regarding an appropriate threshold of foreign language content before submitting that issue to the court.
6. The receiving party may request a specific document or a specific category of documents produced in a particular format (*e.g.*, as TIFF images) to be produced in native format or an alternate format upon good cause shown.

The producing party shall extract and produce the agreed upon metadata, as described below in Section K, from the original Native file, and provide the hash value for the original file.

**1. Native Productions.** All native productions shall be produced in such a way so as to identify and/or maintain the file structure of the data being produced. This shall be accomplished by providing a load file identifying the agreed upon metadata in Section K. Said load files shall be consistent with industry standard Concordance load files such as .DAT files and .OPT files. The file name for documents produced natively shall contain the documents' confidentiality designation, if any. If documents that the parties have agreed to produce in native format per this protocol are to be used at depositions or attached to pleadings or papers filed with the court, the party offering the native document must identify the document as native format and provide the slip sheet provided by the producing party for each native file. A slipsheet shall be produced for all native files at the time of their production which will contain the following fields: the beginning and ending Bates numbers, Protective Order designation, filename, filepath, and hash value.

**(a) Physical Documents.** Documents that either (i) were originally generated as ESI but now only exist in physical hard-copy format, or (ii) were originally generated in hard-copy format, shall be produced as single-page TIFF images along with an OCR acquired text file containing page break indications where available and the following



(b) ESI. For ESI produced as TIFF images, native format files shall be converted to single-page TIFF images. Accompanying the TIFF images shall be a multipage text (.TXT) file containing searchable text from the native file, and the metadata discussed in Section K below. Load files for the TIFF images should be created and produced together with their associated TIFF images to facilitate the use of the produced files by a document management or litigation support database system. The parties shall meet and confer to the extent reasonably necessary to determine how such production can be made reasonably

usable for the requesting party, and to facilitate the import and use of the produced materials with commercially available document management or litigation support software.

**C. Duplicates.** The Parties shall not produce exact duplicates of electronic documents stored in different locations. Each Party must globally de-duplicate identical ESI vertically within each custodian and horizontally across custodians, which will be determined as follows:

(i) Electronic files will be de-duplicated based upon calculated MD5 Hash values for the entire file contents. File contents only will be used for MD5 Hash value calculation and will not include operating system metadata (e.g., filename, file dates) values.

(ii) Messaging files will be de-duplicated based upon MD5 Hash values for the message family, including parent object and attachments. The fields used to create the unique value for each message will be based on the standard settings in commonly available commercial processing software (e.g., iPro, Nuix, or Relativity). Messaging materials will be de-duplicated at a family level, including message and attachment(s).

(iii) When a party de-duplicates identical ESI, the party shall provide custodian associations including fields BATES NUMBER, DUPLICATE CUSTODIANS, and DUPLICATE CUSTODIAN\_SOURCEFILEPATHS as listed in Section K. An updated overlay file containing updated duplicate custodian information shall be produced after every production, if necessary, to account for rolling productions.

(iv) Email Threading. Email threads are email communications that contain prior or lesser-included email communications that also may exist separately in the

party's electronic files. A most inclusive email thread is one that contains all of the prior or lesser-included emails, including attachments, for that branch of the email thread. The producing party need only produce the most inclusive email thread and all attachments from all prior or lesser included email threads, but shall, upon request, also produce all prior or lesser included e-mail threads that were collected and identified as responsive and not privileged. The procedure for logging privileged emails is addressed in Section V below.

**D. Color.** Documents containing color (for example, graphs, pictures, or color marketing materials) will be produced in color for each such document. This provision does not apply to documents solely containing color logos, signatures, watermarks, or letterhead.

**E. System Files.** Electronic file collection will be "De-NISTed," removing commercially available operating system and application files contained on the National Institute of Standards and Technology ("NIST") file list. Identification of NIST list matches will be through MD5 Hash values.

**F. Time Zone/Dynamic Fields.** When processing ESI, Coordinated Universal Time (UTC) should be selected as the time zone. Documents with dynamic fields for file names, dates, and times will be processed to show the field code (e.g., "[FILENAME]" or "[AUTODATE],") rather than the values for such fields existing at the time the file is processed.

**G. Archive File Types.** Archive file types (e.g., .zip, .rar) shall be uncompressed for processing. Each non-privileged, responsive file contained within an archive file shall be produced. If the archive file is itself an attachment, that parent/child relationship shall also be preserved. The filename of the archive file shall appear as part of the SourceFilePath metadata field.

**I. Production Media.** Documents shall be produced on CD-ROM, DVD, external hard drive (with standard PC compatible interface), such other readily accessible computer or electronic media as the parties may hereafter agree upon, or electronically via secure file share/FTP (the "Production Media"). Each production shall be accompanied by a Production Communication

(letter or email) including: (1) the name of the producing party; (2) a legend identifying the production as being made for Valsartan MDL Cases; (3) the production date; (4) a unique production volume, starting with the same Bates prefix used on the documents being produced; (5) the Bates number range of the materials contained on such Production Media item, and (6) information identifying the production as a replacement, if applicable. Each production shall also contain a Production Index, which shall be updated with each production, setting forth the (1) dates, (2) Bates number range, (3) source(s) (*e.g.*, custodians, database) of the materials contained in each production volume, (4) whether the production was a replacement production, and (5) the date of withdrawal of protected designation and replacement with no protected designation for applicable Bates ranges, if applicable.

**J. Electronic Text Files.** For each document, a single text file shall be provided along with the native or image files and metadata. The text file name shall be the same as the Bates number of the first page of the document. File names shall not have any special characters or embedded spaces. Electronic text shall be extracted directly from the native electronic file unless the document was redacted, an image file, or a physical file. In these instances, a text file created using OCR will be produced in lieu of extracted text. Text files will include page numbers and page breaks corresponding to the page numbers and page breaks in the original document to the extent technologically feasible. Where feasible, if the producing party has access to extracted text from electronic document files, the receiving party shall receive extracted text as well, instead of OCR text generated from an image file. If such conversion is unduly burdensome or renders the document unusable, the producing party may produce the document in any reasonably usable form as may be agreed upon by the respective parties. The parties shall meet and confer regarding any disputes relating to the availability of extracted text.

**K. Metadata.** The list attached as Appendix 2 identifies the metadata fields and manually generated fields that will be included in load files, to the extent applicable and available, provided, however, that parties do not need to create metadata fields that are not automatically extracted or extractable by their producing software. Metadata fields may be suppressed or redacted if privileged or information protected by the Health Insurance Portability and Accountability Act of 1996 (“HIPAA”) or other information protected by law. Parties agree to meet and confer if additional fields of metadata are needed.

**L. Parent-Child Relationships.** The parent-child relationship between attachments, enclosures, embedded files, and/or exhibits to any parent document shall be preserved. The child-document should be consecutively produced immediately after the parent-document, unless excluded as a duplicate, in which case a link to the non-duplicate document shall be provided. All parent and child documents shall be produced if any of the parent or child documents are relevant, provided, however, that non-responsive Excel spreadsheets attached to relevant emails need not be produced and may be slipsheeted (as described more fully in Section V.B below). If the parent or child is privileged, and would otherwise be subject to production, a slip sheet shall be inserted in the production and shall be noted in the metadata as a privileged withheld document.

**M. Structured data.** To the extent a response to a non-objectionable discovery request requires or may be completed by production of discoverable electronic information contained in a database, it shall be produced in its native format unless not practicable, in which case the producing party shall notify the requesting party and the parties shall promptly meet to attempt to resolve the issue. If the producing party asserts that production of a database in native is not practicable, the producing party may advise of a preference to (1) produce existing reports or reports readily able to be generated from the database that are reasonably responsive to the

discovery requests, or (2) design queries in order to produce an extract from the database of relevant and responsive data in a reasonably usable and exportable electronic file (e.g., Excel, Access, CSV or Microsoft SQL server database format). Any production of a database not in its native format should be fully parsable, to the extent practicable. For enterprise database systems from which data will be produced, the parties will meet and confer as to the available data in the system, the data to be exported, and the format of production on a case by case basis. The producing party agrees to disclose the search parameters used to design the reports and/or queries prior to the production. Upon review of any such parameters and productions from structured databases, the requesting party may make reasonable requests for additional information, such as database dictionaries, to explain the database schema, fields, codes, abbreviations, and different report formats. The parties shall meet and confer regarding any disputes for such requests.

**N. Data Security.** The parties will ensure that their respective counsel and any third-party vendors hosting data produced in this litigation shall maintain the appropriate level of data security in an effort to prevent a security breach of protected or proprietary information, and pursuant to the Protective Order.

#### **IV. OBJECTIONS TO ESI PRODUCTION**

**A.** If any formatting requirements or other specifications agreed to in this protocol are later determined by the producing party to be not feasible, unduly burdensome or unreasonably costly, or non-proportional to the needs of this case, the parties will promptly meet and confer in order to attempt to agree upon appropriate modifications to the protocol before presenting any such issue to the court.

**B.** Notwithstanding anything contained herein to the contrary, a producing party shall not intentionally produce ESI in a format not requested or agreed to by the requesting party. If an

issue arises, the issue will only be submitted to the court after reasonable efforts by the parties to reach agreement by the meet and confer process in accordance with the Federal Rules of Civil Procedure and Local Civil Rules of this court.

C. If any producing party believes, in good faith, that they cannot or should not be required to produce documents in conformity with this ESI Protocol due to technical or financial reasons, that party shall, at least 14 days prior to the date any production is due, inform the requesting party of the issue, the reason(s) they cannot produce in conformity with this ESI Protocol, and the alternative production method proposed by that party. The requesting and producing parties shall thereafter meet and confer and attempt to resolve the issue. If the parties cannot reach agreement within 14 days of the original notice or such shorter time as required under the circumstances, the issue shall be brought to the Court's attention for resolution at the next discovery conference, or earlier if required under the circumstances.

D. This protocol shall not apply to the production of documents required by the Dismissal Protocol by defendants who are permitted to seek and are seeking dismissal from the action pursuant to the Dismissal Protocol unless agreed to between the parties or ordered by the court based upon a showing of good cause, nor shall it apply to the production of documents by plaintiffs in conjunction with their Plaintiff Profile Forms and Plaintiff Fact Sheets. If, during negotiation of the Plaintiff Fact Sheet, any party believes that this protocol should apply to documents produced in conjunction with the Plaintiff Fact Sheet by any category of plaintiff, the parties shall meet and confer, and if unable to resolve the issue, the contesting party may apply to the court for an order requiring production in conformity with this protocol upon a showing of good cause.

## **V. PRIVILEGE AND WORK PRODUCT CLAIMS AND REDACTIONS**



**A. Privileged Information and Attorney Work Product.** The parties agree that, for documents withheld from production on the basis of attorney-client privilege or the work product doctrine and/or any other applicable privilege, the producing party will provide an updated summary log with each production of privileged material, in an electronically searchable format (e.g., Excel), containing, for each document claimed as privileged, and as required by applicable law, including Fed. R. Civ. P. 26(b)(5)(A), a description of “the nature of the documents, communications, or tangible things not produced or disclosed—and do so in a manner that, without revealing information itself privileged or protected, will enable other parties to assess the claim,” as well as an export of the metadata fields listed below (to the extent such information exists and has not been suppressed or redacted for privilege):

- FIRSTBATES
- LASTBATES
- BEGATTACH
- ENDATTACH
- PARENT\_BATES
- CHILD\_BATES
- ATTACHMENT COUNT
- CUSTODIAN (in the format LASTNAME, FIRSTNAME, MIDDLENAME)
- DUPLICATE\_CUSTODIANS (in the format LASTNAME, FIRSTNAME, MIDDLENAME)
- FROM
- TO
- CC

- BCC
- SUBJECT
- DATE\_SENT
- DATE\_RECEIVED
- FILENAME
- AUTHOR
- DATE\_CREATED
- FILETYPE
- PRIV\_CLAIM (the basis for the claim of privilege, e.g., sent to/from outside counsel)
- STATUS OF PRIVILEGE DESIGNATION (e.g., attorney-client, attorney-work-product)

A party need only log the topmost e-mail in a thread so long as the description of the subject matter includes enough information sufficient to demonstrate the privilege and as long as all lesser-contained e-mails are also independently privileged. If all lesser-contained e-mails are not independently privileged, the topmost e-mail thread must be redacted rather than withheld. To the extent a party seeks to use categorical logs in lieu of providing the information above, the producing party will initiate a meet and confer with the requesting party, and if unable to reach agreement, move the Court to allow such logs upon a showing of good cause. However, documents comprising attorney client communications and/or attorney work product relating solely to the defense of this litigation, and solely between outside counsel or outside counsel and in-house counsel, and dated after the start of the litigation, need not be included on a privilege log. If a producing party believes that documents dated after the start of the litigation for a

particular custodian should not be included on a privilege log, they shall meet and confer with the requesting party, and if unable to reach agreement, move the Court to allow such omission upon a showing of good cause.

If the requesting party requires further information, it shall explain in writing the need for such information and identify, by Bates number or other unique identifier, each document for which it seeks this information. Within thirty (30) days of such a request, the producing party must either (i) provide the requested information or (ii) challenge the request. Either party may seek to expedite this time period upon good cause shown. If a party challenges a request for further information, the parties shall meet and confer to try to reach a mutually agreeable solution. If they cannot agree, the parties must request a conference with the court before any motions may be filed. Withdrawal of privilege designations shall be noted on the summary log, indicating the fact and date of withdrawal of the privilege designation.

The inadvertent production of privileged or protected documents or information shall be governed by Federal Rule of Evidence 502(d) and the Confidentiality and Protective Order entered by the court in this litigation. Consistent with the Confidentiality and Protective Order, and as discussed more fully therein, inadvertently produced privileged or protected documents shall be returned to the producing party or destroyed, upon request, and such returned or destroyed material shall be deleted from any litigation-support or other database, except as set forth in the Confidentiality and Protective Order.

**B. Redactions.** A producing party may redact ESI that the producing party believes is subject to attorney-client privilege, is protected work product, or is otherwise legally prohibited from disclosure. A producing party may also redact information contained in a document (or, in the case of an Excel spreadsheet attached to a broader family group of documents, withhold a

document in its entirety) that is non-responsive because it does not concern the product(s) at issue nor is relevant to any party's claims or defenses, provided that any such information meets the criteria for designation as "Confidential Information" or "Restricted Confidential Information," and the non-responsive information is segregable from the responsive information contained in the document such that redaction of the non-responsive information will not impede the receiving party's ability to interpret the responsive information contained in the document.

The producing party shall mark each redaction with the bases for the redaction, and a redacted document's privilege designation will be made visible on the face of the document. Where a producing party redacts information that is non-responsive because it does not concern the product(s) at issue nor is relevant to any party's claims or defenses, the producing party shall label the redaction "Redacted – Other Product(s)." Where a document is withheld in its entirety because the document does not concern the product(s) at issue nor any party's claims or defenses, a slipsheet will be provided labeled with the Bates number and "Withheld – Other Product(s)." Even where a document is withheld in its entirety under this provision, however, all metadata associated with such document will still be provided.

The parties shall log redacted documents on a privilege log in the same manner as that for a privileged document.

## **VI. DESTRUCTION AND RETURN OF ESI**

When a producing party believes, in good faith, that this litigation and all related litigations have fully and finally terminated, the producing party shall request the receiving party to return or destroy all confidential discovery materials produced. If there is a dispute as to whether or not such return or destruction is appropriate at the time of the request, the parties shall meet and confer, and if they cannot resolve the issue, bring the issue before the court. Within sixty (60)

days after agreement to return or destroy, or the entry of an order authorizing return or destruction not subject to further appeal, all confidential discovery materials produced by the requesting party must be either destroyed or returned to the producing party. If destroyed, the receiving party must produce an affidavit certifying the destruction to the producing party within thirty (30) days of destruction.

#### **VII. AMENDMENTS TO THE ESI PROTOCOL**

No modifications or amendments to the ESI Protocol shall be made except by agreement by all signatories or in conformity with the Federal Rules of Civil Procedure and the Local Civil Rules of this court.

IT IS SO ORDERED.

Dated: June 19, 2019

  
\_\_\_\_\_  
HON. JOEL SCHNEIDER  
UNITED STATES MAGISTRATE JUDGE

## APPENDIX 1: FILE FORMATS

### A. Load Files

- Every electronic document referenced in a production load file shall have all corresponding file, text, and data logically grouped together in a directory structure with a common key to properly load the data. This will include a native link (path to the native file as included in the production e.g., d:\PROD001\natives\ABC00015.xls) for all files produced in native format.
- The name of the image load file shall mirror the name of the delivery volume, and should have an .lfp, .opt or .dii extension (e.g., ABC001.lfp).
- The volume names shall be consecutive (e.g., ABC001, ABC002, et. seq.).
- The producing party will provide document level text files (extracted text or OCR) to accompany the production. Each document's electronic image will convey the same information and image as the original document.
- The image load file shall contain one row per file produced.
- Every image in the delivery volume shall be contained in the load file.
- The file key shall be named the same as the Bates number of the page.
- Load files shall not span across media (e.g., CDs, DVDs, Hard Drives, etc.), i.e., a separate volume shall be created for each piece of media delivered.

### B. Metadata Load Files

- The metadata load file shall use the following delimiters:
  - Column Delimiter: Pilcrow - ¶ (ASCII :020)
  - Text Qualifier: Thorn - þ (ASCII :254)
  - New line: Registered sign - ® (ASCII 174)
- The first record shall contain the field names in the order of the data set forth in Section K above.
- Metadata fields that are not applicable to a document shall be filled with two adjacent text delimiters.
- All date fields shall be produced in "mm/dd/yyyy hh:mm AM" format.
- A carriage-return line-feed shall be used to indicate the start of the next record.

- Load files shall not span across media (e.g., CDs, DVDs, Hard Drives, etc.); a separate volume shall be created for each piece of media delivered.
- The name of the metadata load file shall mirror the name of the delivery volume, and shall have a .dat, .csv or .txt extension (i.e., ABC001.dat).
- The volume names shall be consecutive (e.g., ABC001, ABC002, et. seq.).
- The .dat and .txt files shall be encoded in UTF-8 or UTF-16 coding as appropriate, preserving the integrity and full character set in use for any Japanese or other non-English text, and shall be uniform across all productions.

#### C. Scanned Hardcopy Documents

- In scanning hardcopy documents, multiple distinct documents should not be merged into a single record, and single documents should not be split into multiple records (*i.e.*, hard copy documents should be logically unitized).
- OCR should be performed on a document level and provided in document-level \*.txt files named to match the production number of the first page of the document to which the OCR text corresponds. OCR text should not be delivered in the data load file or any other delimited text file. The text file shall contain image keys/Bates numbers for each page.
- In the case of an organized compilation of separate hardcopy documents—for example, a binder containing several separate documents behind numbered tabs—the document behind each tab should be scanned separately, but the relationship among the documents in the binder should be reflected in proper coding of the attachment fields.

## APPENDIX 2: METADATA FIELDS

Field	Description
FIRSTBATES	Bates number of first page of document (for imaged collections only) OR document-level ID number assigned on a file level (for native-only productions)
LASTBATES	Bates number of last page of document (for imaged collections only)
BEGATTACH	Bates number of the first page of the first document in the range.
ENDATTACH	Bates number of the last page of the last attachment in the range.
MAPI ATTACHMENT COUNT	Count of email attachments from MAPI field
ATTACHMENT_COUNT	Count of email attachments actually produced by production software where technically feasible
CUSTODIAN (in the format LASTNAME, FIRSTNAME, MIDDLENAME)	Person, organization, or department from whom materials were collected.
DUPLICATE_CUSTODIANS (in the format LASTNAME, FIRSTNAME, MIDDLENAME)	All custodians who had a copy of the document in their data collection.
SOURCEFILEPATH	Full file path of the original native document
DUPLICATE_CUSTODIAN_SOURCEFILEPATHS	Full file path of the original native document for duplicate custodians
FROM	Email "From" data
TO	Email "To" data (semi-colon delimited, if multiple entries)
CC	Email "CC" data (semi-colon delimited, if multiple entries)



BCC	Email "BCC" data (semi-colon delimited, if multiple entries)
SUBJECT	Email Subject
DATE_SENT (UTC TIME ZONE) (for emails)	Email Date Sent
DATE_RECEIVED (UTC TIME ZONE) (for emails)	Email date received
FILENAME	System long file name
FILE_EXTEN	System file extension
AUTHOR	Application-stored user/author metadata
DATE_CREATED	System-created document creation date
DATE_MOD	System-created document last-modified date
DATE_LASTACCESS	System-created document last-accessed date
LAST_PRINTED_DATE	System-created document last-printed date
FILE_SIZE	Numeric file size in KB
FILETYPE	File Type as determined by producing software (e.g. Word, Excel, PowerPoint)
RECORDTYPE (Paper/Electronic)	Whether document was a paper document which was scanned at the time of production, or was an electronic document at the time of production
INTMSGID	Message ID value extracted from Microsoft email data
HASH_VALUE	MD5 value of the original electronic file

TEXTLINK	Full file path of document-level text file on production media
NATIVELINK	Full file path of native document on production media
PAGE_COUNT	Calculated page count of document
EMAIL_IMPORTANCE (YES if flag present, blank if not)	Message importance flag from Outlook email
REPLACEMENT_DATE	Date on which the document was produced as a replacement document, if applicable
PROD_VOLUME	Production volume name
FOREIGN_LANGUAGE (YES if flag present, blank if not)	System Detected Foreign Language
CONFIDENTIALITY DESIGNATION	Populated with the applicable confidentiality designation (CONFIDENTIAL, HIGHLY CONFIDENTIAL, etc.) associated with the produced file
WITHHELD_PRIVILEGED (YES if flag present, blank if not)	Attachment withheld as privileged and slipsheeted
REDACTED (YES if flag present, blank if not)	Document contains redactions
REDACTION_TYPE	The basis for the redaction (e.g., attorney-client, other product)
TITLE (EDOCS)	Title as extracted by producing software
SUBJECT (EDOCS)	Subject as extracted by producing software
CONVERSATION_INDEX	Conversation index extracted from Microsoft email data
APPLICATIONNAME	Application-stored program name if determinable by production software

CATEGORIES	Application-stored category value
COMMENTS	Application-stored comments metadata
DELIVRECEIPT (YES if flag present, blank if not)	Message-delivered flag from Outlook email
HASCOMMENTS (YES if flag present, blank if not)	Flag whether document has comments
HASTRACKCHANGES (YES if flag present, blank if not)	Flag whether document has track changes
HIDENTEXT (YES if flag present, blank if not)	Flag whether document has hidden text
KEYWORDS	Application-stored keywords metadata
LANGUAGE	Language as determined by production software (e.g. German, French)
LASTAUTHOR	Last Edited By field in metadata
ORIGFILE	Original file name from native PDF file, if available in producing software
ORIGEXT	Original file extension from native PDF file, if available in producing software
READRECEIPT (YES if flag present, blank if not)	Message-read flag from Outlook email
REVISION	Application-stored revision number value

27